Erra and Gabpa/b specify PGC-1a-dependent OXPHOS gene expression that is altered in diabetic muscle

Vamsi K. Mootha[1*], Christoph Handschin[2*], Dan Arlow[1], Xiaohui Xie[1], Julie St. Pierre[2], Smita Sihag[1], Wenli Yang[2], David Altshuler[1], Pere Puigserver[2,4], Patricia J. Willy[3], Ira G. Shulman[3], Richard A. Heyman[3], Eric S. Lander[1], and Bruce M. Spiegelman[2]

[1]MIT/Broad Institute, One Kendall Square, Cambridge, Massachusetts 02139; [2]Dana-Farber Cancer Institute and Department of Cell Biology, Harvard Medical School, One Jimmy Fund Way, Boston, Massachusetts 02115; [3]X-Ceptor Therapeutics Inc., 4757 Nexus Center Drive, San Diego, California 92121

*equal contribution

Published in Proc Natl Acad Sci U S A. 2004 Apr 27;101(17):6570-5. PMID: 15100410. doi: 10.1073/pnas.0401401101

Errα and Gabpa/b specify PGC-1α-dependent OXPHOS gene expression that is altered in diabetic muscle

Vamsi K. Mootha[1*], Christoph Handschin[2*], Dan Arlow[1], Xiaohui Xie[1], Julie St. Pierre[2], Smita Sihag[1], Wenli Yang[2], David Altshuler[1], Pere Puigserver[2,4], Patricia J. Willy[3], Ira G. Shulman[3], Richard A. Heyman[3], Eric S. Lander[1], and Bruce M. Spiegelman[2]

[1]MIT/Broad Institute, One Kendall Square, Cambridge, Massachusetts 02139; [2]Dana-Farber Cancer Institute and Department of Cell Biology, Harvard Medical School, One Jimmy Fund Way, Boston, Massachusetts 02115; [3]X-Ceptor Therapeutics Inc., 4757 Nexus Center Drive, San Diego, California 92121

*These authors contributed equally to this work

Correspondence should be addressed to:

Vamsi K. Mootha, MIT/Broad Institute, One Kendall Square, Bldg. 300, Cambridge, MA 02139-1561, Phone: 617 252-1672, Fax: 617 252-1902, Email: vmootha@broad.mit.edu

or Bruce M. Spiegelman, Dana-Farber Cancer Institute, Department of Cancer Biology, Smith Building SM958, One Jimmy Fund Way, Boston, MA 02115, Phone: 617 632-3567, Fax: 617 632-4655, Email: bruce_spiegelman@dfci.harvard.edu

[4] Current address: Department of Cell Biology, Johns Hopkins University Medical

School, Baltimore, Maryland 21205

Recent studies have shown that genes involved in oxidative phosphorylation (OXPHOS) exhibit reduced expression in skeletal muscle of diabetic and prediabetic humans. Moreover, these changes may be mediated by the transcriptional co-activator peroxisome proliferator-activated receptor g co-activator-1a (PGC-1a) [1,2]. By combining PGC-1a-induced genome-wide transcriptional profiles with a computational strategy to detect *cis*-regulatory motifs, we identified estrogen receptor-related receptor a (Erra) and GA-binding protein a (Gabpa) as key transcription factors regulating the OXPHOS pathway. Interestingly, the genes encoding these two transcription factors are themselves PGC-1a-inducible and contain variants of both motifs near their promoters. Cellular assays confirmed that Erra and Gabpa partner with PGC-1a in muscle to form a double-positive feedback loop that drives the expression of many OXPHOS genes. By using a synthetic inhibitor of Erra, we demonstrated its key role in PGC-1a-mediated effects on gene regulation and cellular respiration. These results illustrate the dissection of gene regulatory networks in a complex mammalian system, elucidate the mechanism of PGC-1a action in the OXPHOS pathway, and suggest that Erra agonists may ameliorate insulin-resistance in individuals with type 2 diabetes mellitus.

Peroxisome proliferator-activated receptor-γ co-activator 1-α and 1-β (PGC-1-α and -β) responsive genes involved in oxidative phosphorylation show reduced expression in muscle from patients with type 2 diabetes mellitus. These changes are also observed in healthy first degree relatives of diabetics as well as in individuals with impaired glucose tolerance, suggesting that these defects are causally linked to the development of insulin resistance. Combined with recent morphologic and functional studies [3-5], these reports indicate that a PGC-1-mediated dysregulation in mitochondrial biogenesis is important in the development of human diabetes.

PGC-1-α and -β are transcriptional co-activators involved in the regulation of several key metabolic processes, including mitochondrial biogenesis, adaptive thermogenesis, skeletal muscle fiber-type switching, and gluconeogenesis [6-12]. PGC-1α is capable of co-activating nearly all known nuclear receptors as well as many other transcription factors [13]. A pressing question now is to identify the transcription factors that specify the changes observed in muscle: these transcription factors may be sites of genetic variation or misregulation and may represent important therapeutic targets.

Systematic identification of transcription factors involved in biological processes in mammals remains a largely unsolved problem [14]. One promising approach relates genome-wide expression profiles to promoter sequences to discover influential *cis*-motifs [15-18]. Such strategies have yielded impressive results in simple organisms such as yeast, but it has been challenging to extend these algorithms to mammalian genomes, where intergenic regions are large, annotation of gene structure is imperfect, and DNA sequence can be highly repetitive. Most of these methods seek motifs by comparison to a fixed background model of nucleotide composition (which fails to

represent the fluctuations seen in large genomes) or by comparison between two sets of genes (which is likely to capture only very sharp differences). Further, many of these methods assume that the expression data are normally distributed, which may not always be true.

To overcome some of these obstacles, we devised a simple, nonparametric strategy for identifying motifs associated with differential expression (motifADE) (Fig. 1a). The algorithm involves three steps: (i) ranking genes based on differential expression between two conditions; (ii) given a candidate motif M, identifying the subset S_M of genes whose promoter regions contains the motif M; and (iii) testing via a nonparametric statistic (see Methods) if the genes in S_M tend to appear toward the top (or bottom) of the ranked list (indicating association) or are randomly distributed on the list. motifADE may be applied to a specific candidate motif of interest or to the list of all possible motifs of a given size (in which case the significance level should be adjusted to reflect multiple hypothesis testing). By using a nonparametric scoring procedure, we do not make assumptions about the distribution of the expression data. Furthermore, by considering the entire rank ordered list, the promoters of lower rank implicitly provide a background of DNA composition for comparison, and there is no need to demarcate or cluster a group of genes as being "responsive." The method is quite flexible, and can test any set of motifs using a traditional promoter database or even a database of promoters that have been masked based on evolutionary conservation (see Methods).

To identify motifs related to PGC-1α action, we infected mouse C2C12 muscle cells with an adenovirus expressing PGC-1α and obtained gene expression profiles for 12,488 genes at 0, 1, 2 , and 3 days following infection. We found 649 genes that were

induced at least 1.5-fold at day 3. As expected, these were enriched for genes involved in carbohydrate metabolism and the mitochondrion [6,7,11,12,19] (Fig. 1c). Somewhat surprisingly, many genes involved with protein synthesis (GO terms: protein biosynthesis, mitochondrial ribosome and ribosome) are also induced.

We then applied motifADE to study the 5034 genes for which there is reliable information about the location of the transcriptional start site (TSS) in the mouse genome (see Methods). For each gene, the target region was defined to be a 2 Kb region centered on the TSS. We then tested all possible *n*-mers ranging in size from n= 6 to n=9 nucleotides (a total of 348,160 motifs) for association with differential expression on each of the three days of the timecourse.

A total of 20 motifs achieved high statistical significance ($p<0.001$, corrected for multiple hypothesis testing) and these were almost exclusively related to two distinct motifs (see Table 1 as well as the longer list in Supplemental Table S1). The first motif, 5'-TGACCTTG-3' was significant on days 1, 2, and 3 (adjusted $P=2.1\text{x}10^{-6}$, $2.9\text{x}10^{-9}$, and $7.7\text{x}10^{-7}$, respectively). It corresponds to the published binding site for the orphan nuclear receptor Errα [20], which is known to be capable of being co-activated by PGC-1-α and -β [21-23]. The *Errα* gene is known to be involved in metabolic processes, based on studies showing that knockout mice have reduced body weight and peripheral fat tissue, as well as altered expression of genes involved in metabolic pathways [24]. The second motif is 5'-CTTCCG-3' (adjusted $p=8.9\text{x}10^{-9}$), which is the top scoring motif on day 3. It corresponds to the published binding site for Gabpa [25], which complexes with Gabpb [26] to form the heterodimer nuclear respiratory factor-2 (NRF-2) [27], a factor known to regulate the expression of some OXPHOS genes [28]. Interestingly, the reverse

complements of these motifs did not score well, suggesting a preference for the orientation of these motifs relative to the TSS, and some occurrences of the motifs occurred downstream of the TSS. While each of these motifs is individually associated with PGC-1α, our analyses suggest that a gene having both motifs typically ranks higher on the list of differentially expressed genes (Supplemental Fig. S1), suggesting that the two motifs might have an additive or synergistic effect.

Next, we repeated the motifADE analysis using a "masked" promoter database (see results in Supplemental Table S2). We still considered the 2000 bp centered on the TSS, but masked these promoters so that we only considered those nucleotides aligned and conserved between mouse and human (see Methods). Still, the top ranking motifs on days 1 and 3 were related to Errα (day 1, P=6.5x10^{-7}; day 3 P=4x10^{-7}) and to Gabpa (day 3 P=5.2x10^{-5}), providing additional support for these motifs.

The Errα and Gabpa motifs are particularly enriched in the OXPHOS-CR genes, which have been shown to be reduced in human diabetes [1,2]. Whereas the top scoring Errα motif (5'-TGACCTTG-3' or its reverse complement) only occurs in 12% of the 5034 promoters of the database, in 29% of the PGC-responsive genes (*i.e.*, those genes induced at least 1.5 fold on day 3), and in 27% of the mitochondrial genes, they are found in 52% of the OXPHOS-CR genes (significance of enrichment, P=1x10^{-4}). About one-half of these sites are perfectly conserved in the syntenic region in human (data not shown). The top scoring Gabpa binding sites (5'-CTTCCG-3' or its reverse complement) are much more common (62% of all promoters of the database and in 79% of the PGC-responsive genes), but they, too, show significant enrichment in the OXPHOS-CR genes (89%, P=0.02).

The results suggest that Errα and Gabpa may be key transcriptional factors mediating PGC-1α action in muscle. In this connection, it is notable that based on the microarray data, both Errα and Gabpa are themselves induced ~2-fold (P<0.01) on day 1 following expression PGC-1α. The induction of these factors has been observed in previous studies [7,21]. Moreover, careful analysis of the *Erra* and *Gabpa* genes suggest that each contain potential binding sites for both transcription factors within the vicinity of their promoters. The *Erra* gene has the Errα motif as well as a conserved variant of the Gabpa binding site [25] upstream of the TSS, while the *Gabpa* gene has an Errα site upstream of the TSS and a conserved variant of the Gabpa binding site in its first intron (see Fig. 2a and Methods). These results raise the possibility that Errα and Gabpa may regulate their own and each other's expression.

Taken together, the systematic analysis of the transcriptional program driven by PGC-1α in skeletal muscle suggests a model in which increases in PGC-1α protein levels (induced, for example, by exercise [29,30]) results in increased transcriptional activity of Gabpa and Errα on their own promoters, leading to a stable increase in the expression of these two factors via a double positive-feedback loop (Fig. 2). These two factors, perhaps in combination with PGC-1α, are then crucial in the induction of downstream target genes, many of which have binding sites for these motifs (Fig. 2). Such a circuit may serve as a regulatory switch, analogous to a feed-forward loop that plays a key role in the early stages of endomesodermal development in sea urchin [31].

We experimentally tested this model of gene regulation. To determine whether Errα and Gabpa are indeed key transcriptional partners of PGC-1α, we tested whether the Errα and Gabpa proteins are co-activated by PGC-1α on the *Erra* and *Gabpa*

promoters. We cloned appropriate portions of the promoters (the upstream region of *Erra* containing both sites, the upstream region of *Gabpa* containing the Errα site and the first intron of *Gabpa* containing the Gabpa site) into expression vectors and tested their ability to drive expression of a reporter gene, in the presence and absence of PGC-1α. The results confirm the predicted co-activation between PGC-1α and the two transcription factors in muscle cells. Moreover, the results for the *Erra* promoter fragment containing both sites show that the transcription factors act additively or even synergistically.

We next sought to explore the functional role of Errα in executing the PGC-1α-mediated response, by using a recently developed, selective Errα inverse agonist, called XCT790, referred to below as the inhibitor. This compound does not significantly inhibit or activate other nuclear receptors at doses near its IC_{50} (583 nM) for Errα, although higher concentrations (above 3.3 μM) have been found to weakly activate PPARγ (R.A. Heyman, personal communication).

The inhibitor was first used to probe the role of Errα in activation of the Errα and the Gabpa, by repeating the experiments above in the presence or absence of the inhibitor at a dose slightly above its IC_{50}. The inhibitor markedly reduced activation of the *Erra* promoter by Errα and PGC-1α, but had no effect on Gabpa/b-driven reporter gene expression (Fig. 4a). The effect of the Errα inhibitor resembles the effect of site-directed mutagenesis of the Errα binding site on the *Erra* promoter (Fig. 4b).

The role of Errα was investigated in the regulation of endogenous PGC-1α target genes. PGC-1α was expressed in C2C12 myotubes, treated with inhibitor or vehicle,

and measured the expression level of specific genes at day 1 by using RT-PCR. We studied five genes that contain variants of the Errα motif, including two of the OXPHOS-CR genes, Gabpa, and Errα, as well as medium chain acyl-CoA dehydrogenase (MCAD), a published target of Errα [22,32]. The expression of all of these genes is induced by PGC-1α, and this induction is diminished by the inhibitor (Fig. 4c). To test the specificity of the response, we also evaluated two genes that are induced by PGC-1α but yet lack obvious Errα binding sites. These two genes, pyruvate dehydrogenase β and the inner mitochondrial membrane protein Immt, were not affected by the inhibitor (Fig. 4c). Together, these results suggest that a subset of PGC-1α genes are regulated by Errα and that these correspond to the target genes containing an Errα motif. The fact that the inhibition of these genes by the ERRα inverse agonist was not complete suggests that other factors might also be involved in this regulation. To ensure that the effects of the Errα inhibitor were not due to activation of PPARγ, control experiments were performed using the synthetic PPARγ agonist rosiglitazone and confirmed that it had no effect on the expression of these genes (Supplemental Fig. S2).

The role of Errα in the PGC-1α-mediated physiologic response was characterized using functional assays of mitochondrial respiration. The Errα inhibitor potently diminishes PGC-1α induction of total cellular respiration and also significantly reduces PGC-1α triggered increases in uncoupled respiration (Fig. 5). The data suggest that inhibition of Errα elicits certain aspects of a 'diabetic phenotype' in cultured muscle cells.

We extended the analysis to cells derived from liver, where PGC-1α has been shown

to induce genes of the fasting response, including those involved in gluconeogenesis and β-oxidation of fatty acids [8,33,34]. These responses are generally elevated in poorly controlled diabetics. As expected, MCAD and two key genes related to gluconeogenesis (glucose-6-phosphatase and phosphoenol-pyruvate carboxykinase) were all induced by PGC-1α. However, the Errα antagonist repressed PGC-1α-induction of MCAD but not the gluconeogenic genes (Fig. 4d).

Finally, we investigated the relationship of nuclear respiratory factor-1 (NRF-1) to ERRα function in muscle cells. NRF-1 has been suggested to be a key transcription factor involved in mitochondrial biogenesis [35] and is induced and co-activated by PGC-1α in this process [7]. On the other hand, transgenic overexpression of NRF-1 in muscle is not sufficient to induce mitochondrial biogenesis [36]. Using the synthetic Errα inverse agonist, we found that induction of NRF-1 by PGC-1α is actually downstream of the initial Errα-mediated events (Supplemental Fig. S3). Interestingly, motifADE analysis of the NRF-1 consensus motif (5'-GCGCAC/TGCGC-3' or reverse complement) yields a score which falls far short of significance in the search over all possible motifs (in which P-values must be corrected for multiple hypothesis testing) but which is significant if considered as a test of single hypothesis (nominal P=0.001 on day 3). Overall, these results suggest that NRF-1 is an important but relatively late mediator of the mitochondrial transcriptional program induced by PGC-1α.

The results above have computational, biological and medical implications. First, the motifADE algorithm provides a simple, nonparametric approach for discovering *cis*-elements by incorporating gene expression. The method makes very few assumptions about the statistical properties of DNA composition or about the distribution of RNA

expression. The method is flexible, and as we have shown, can easily incorporate "masked" or "phylogenetically footprinted" promoters from cross-species comparisons. With additional cross-species comparisons, it should be possible to further reduce the search space[37-39]. Fortunately, in the current study, we were confident in the identity of the transcription factors binding the motifs discovered – in general this may not be the case, and generic, experimental strategies will be needed to systematically determine the occupancy of newly identified motifs. The strategy clearly has some limitations – *e.g.*, a motif may be missed if it lies outside the target promoter region, or if a functional binding site is too degenerate for our motif search strategy.

Second, the analyses above indicate that the effects of PGC-1α on OXPHOS genes in muscle are largely mediated through Errα and Gabpa. Recent studies have shown that PGC-1β can also co-activate Errα [23], and initial application of motifADE to cells expressing PGC-1β suggests that these same factors are important in regulating this program also (data not shown). Together, the data imply a model of gene regulation in which PGC-1α (and likely PGC-1β) initially induces the expression of Errα and Gabpa, via a double positive feedback mechanism (Fig. 2). These transcription factors are then expressed at higher levels and are themselves co-activated by PGC-1 to induce downstream genes such as NRF-1 and members of OXPHOS. Certainly, other transcription factors and regulators, not identified in the current study, are involved in the mitochondrial biogenesis program. Whereas previous studies have shown that PGC-1 interacts with and/or induces 15-20 transcription factors in various physiological settings (including Errα and Gabpa [7,21-23,40]), the present study points to Errα and Gabpa as being especially important early in the timecourse in muscle and provides a

model of how these factors interact in executing the transcriptional program.

Third, the results suggest a potential approach to the treatment of type 2 diabetes. Recent studies in diabetic and pre-diabetic humans have demonstrated that there is a consistent decrease in the expression of genes of oxidative phosphorylation that are responsive to PGC-1α and PGC-1β and that treatments that induce PGC-1α (such as exercise) lead to increased expression of OXPHOS genes and improved insulin sensitivity [1,2,4,5]. On its face, this might argue for developing therapeutic approaches that raise the transcriptional activity of PGC-1. However, PGC-1 activates many different pathways in many tissues and such approaches may suffer from lack of specificity. For example, global transgenic overexpression of PGC-1β in mice results in resistance to obesity induced by a high-fat diet or by a genetic abnormality, though the contribution of PGC-1β expression in muscle has not been explored [23]. On the other hand, a global knockout of *Errα* also causes a leaner phenotype and resistance to high-fat diet-induced obesity [24]. Combined with the current results, this suggests that Errα has different functions in energy homeostasis in different tissues. The identification of the critical roles of Errα and Gabpa in mediating the transcriptional program altered in human diabetic muscle may offer a more specific target. Errα plays a key role in muscle cells but does not appear to affect genes involved in gluconeogenesis in liver cells. This implies that synthetic Errα agonists, particularly those that promote docking of PGC-1α on Errα, may serve to ameliorate insulin resistance in human diabetic muscle without exacerbating hyperglycemia in the liver. Because Errα is an orphan nuclear receptor, it may be an attractive drug target.

Methods

Cell culture and adenoviral infection. Mouse myoblasts (C2C12 cells) were cultured and differentiated into myotubes as previously described [7]. After 3 days of differentiation, they were transduced with an adenovirus expressing either green fluorescence protein or PGC-1α. All experiments were performed in duplicate and flow cytometry performed to ensure nearly 100% and equivalent infection titers.

mRNA isolation, target preparation, and hybridization. Cells were harvested prior to (day 0) as well as 1, 2, and 3 days following viral transduction (all performed in duplicate, yielding 14 samples). RNA was isolated from the cultured cells using the Trizol method. cRNA targets were prepared as previously described [1] and hybridized onto the Affymetrix MG-U74Av2 chip. Hence, the timecourse consists of a total of 14 microarray experiments.

Data scaling. Data were scaled with a mean scaling procedure using the median scan as a reference. The data files were then converted into the .res format for further analysis. Complete timecourse data can be downloaded at http://www-genome.wi.mit.edu/mpg/PGC_motifs/.

Data Clustering, Visualization, and Annotation Enrichment. Data were clustered and visualized using the dChip software package [41]. Normalized gene expression levels were used for data visualization. Mitochondrial genes were defined from a recent report that combined proteomics and genomics to systematically identify mouse

mitochondrial proteins [42]. The GoSurfer tool was used to determine which annotations were enriched groups of genes, using a *P*-value of 0.01 [19].

Identifying Genes Responsive to PGC-1α. For a given day, simple *t*-tests were performed to identify whether a given gene shows difference in expression between PGC-1α and GFP infection. We required a 1.5-fold increase in expression and a nominal *P*-value of 0.05 as arbitrary thresholds for identifying differentially expressed genes.

Promoter Databases. We used the mm3 build of the mouse genome (http://genome.ucsc.edu) and the associated RefSeq annotation tables to construct a database of mouse promoters corresponding to the probe-sets on the Affymetrix MG-U74Av2 chip. The MG-U74Av2 chip contains a total of 12,488 probe-sets. The mm3 RefSeq includes annotations for 13,406 reference sequences. Of the 12,488 probe-sets on the MG-U74Av2 chip, 5034 map 1:1 to a RefSeq that maps uniquely in the mouse genome. Our 'mouse promoter database' consists of the 1000 bases upstream and downstream (2000 BP total) surrounding the annotated transcription start site of these 5034 genes.

We also performed analyses on a 'masked promoter database', consisting of the regions within these 2000 BP that are aligned and conserved between mouse and human. We used the mouse/human BLASTZ alignments (mouse mm4 vs. human hg16) [43] that we downloaded from the UCSC genome browser (http://genome.ucsc.edu/goldenPath/mm4/vsHg16/axtNet/). We only considered the

4249 promoters for which the mouse:human BLASTZ alignment contained at least 100 BP. Finally, we masked the aligned promoters to retain mouse sequence exhibiting at least 70% sequence identity to human within windows of size 10. The median promoter length in the masked database is ~ 1200 BP, compared to a fixed length of 2000 BP in the mouse promoter database.

Motif discovery. To discover motifs associated with PGC-1α responsiveness, we developed a program called motifADE, which a variant of the Gene Set Enrichment Analysis procedure recently described [1]. Comparisons are made for each day between the PGC-1α treated versus GFP control group. All genes (corresponding to genes in the promoter database) on a given day are rank ordered on the basis of difference in expression, based on the signal to noise ratio (SNR) from the microarray experiment. In general, the genes can be ranked by any difference metric. Each gene is then annotated for the presence or absence of a motif in the promoter. We used string matching to search for *n*-mer (where n = 6, 7, 8 or 9) or for selected motifs of interest defined by a regular expression.

To determine whether the differential expression for those genes with a given motif differ from those genes lacking the motif, we used the Mann-Whitney rank sum statistic. Because the masked promoter database consists of promoters of unequal length, a motif could potentially be associated with a gene simply because of promoter length. We corrected the Mann-Whitney statistic by explicitly taking the promoter length into consideration. The null hypothesis of our correction is that a given motif occurs in a promoter with a probability that depends linearly on the promoter length. This

distribution will be uniform for promoters sharing the same length, thereby reducing to the null hypothesis used in the traditional Mann-Whitney statistic.

To estimate the significance of a motif that occurred in *C* promoters with a rank sum of *S,* we compute a *Z*-score = $(S-\mu)/\sigma$, where the mean μ and standard deviation σ is determined from the new null hypothesis. Monte Carlo simulations are used to estimate σ and μ of the rank sum in the null hypothesis by randomly drawing samples of *C* promoters according to the given probability distribution. For each *C*, 1000 samples were drawn and μ and σ of the rank sum estimated from the sample mean and the sample deviation.

The *P*-values for the association between motifs and expression have been subjected to a Bonferroni correction for the 4^n *n*-mers tested, *i.e*, the *P*-values in Table 1 represent global or adjusted *P*-values. We also applied motifADE to the expression data, testing individual, published motifs. The *P*-values reported for the association of these motifs are nominal *P*-values, since these represent single hypotheses. Finally, we note that these *P*-values were computed under a null model of random distribution.

motifADE has been implemented in C and is available upon request (V.K.M.)

Cotransfection assays. The Errα and the Gabpa promoters and the Gabpa intron 1 were amplified from genomic mouse DNA by PCR and subsequently cloned into the pGL3 reporter gene vectors (Promega). Transfections and reporter gene assays were performed as described [44]. Site-directed mutagenesis was done using overlapping primers. Simultaneously with the transfection, cells were treated with vehicle (0.1% DMSO) or 1 μM of the inhibitor, XCT790, for 48 hours. Motifs for Errα and Gabpa were

found at the following positions: Errα site at -585 CCGCAG**TGACCTTGA**GTTTTG -565 and Gabpa site at -41 GGA**GGAA**GCGGAGTA**GGAA**GCA -20 in the Errα promoter, Errα?site at -614 GCGCCG**TGACCTTGG**GCTGCC -594 in the Gabpa promoter and Gabpa sites at +3 GAG**TTGCGG**ACG +14 and +51 TGG**TTCCGG**GGC +62 in intron 1 of the Gabpa gene. These motifs were either directly identified by motifADE or, in the case of the Gabpa binding site in the Errα promoter, found by looking for other Gabpa motifs that have been described [26].

Semiquantitative real-time PCR. Primers for target genes were designed using the PrimerExpress software (ABI Biosystems). Relative gene expression was calculated with the ΔΔCt method using SYBRgreen on a iCycler (Biorad).

Respiration assays. After differentiation into myotubes, C2C12 cells were infected with adenovirus encoding for GFP or PGC-1α and treated with vehicle or XCT790. Two days post-infection, total mitochondrial respiration and uncoupled respiration were measured as published [45].

URLs. The mouse genome sequence, RefSeq annotation tables, and BLASTZ alignments were downloaded from http://genome.ucsc.edu. Mappings from Affymetrix probe-sets to RefSeqs were performed using the NetAffx website (http://www.affymetrix.com). Complete datasets used in this manuscript are available for download at http://www-genome.wi.mit.edu/mpg/PGC_motifs/.

Acknowledgements.

We are grateful to H. Bolouri and N. Patterson for valuable discussions. V.K.M. is supported by the Howard Hughes Medical Institute as a physician postdoctoral fellow. C.H. is supported by a fellowship of the "Schweizerische Stiftung für Medizinisch-Biologische Stipendien", the Swiss Academy of Medical Sciences and the Swiss National Science Foundation. J.S. is a Merck Fellow of the Jane Coffin Childs Memorial Fund for Medical Research. This work was supported in part by National Institute of Health grants to B.M.S.

References.

1. Mootha, V.K. et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-73 (2003).
2. Patti, M.E. et al. Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. *Proc Natl Acad Sci U S A* **100**, 8466-71 (2003).
3. Kelley, D.E., He, J., Menshikova, E.V. & Ritov, V.B. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. *Diabetes* **51**, 2944-50 (2002).
4. Petersen, K.F. et al. Mitochondrial dysfunction in the elderly: possible role in insulin resistance. *Science* **300**, 1140-2 (2003).
5. Sreekumar, R., Halvatsiotis, P., Schimke, J.C. & Nair, K.S. Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. *Diabetes* **51**, 1913-20 (2002).
6. Puigserver, P. et al. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829-39 (1998).
7. Wu, Z. et al. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115-24 (1999).
8. Yoon, J.C. et al. Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131-8 (2001).
9. Yoon, J.C. et al. Suppression of beta cell energy metabolism and insulin release by PGC-1alpha. *Dev Cell* **5**, 73-83 (2003).

10. Lin, J. et al. Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres. *Nature* **418**, 797-801 (2002).

11. Vega, R.B., Huss, J.M. & Kelly, D.P. The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor alpha in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Mol Cell Biol* **20**, 1868-76 (2000).

12. Lehman, J.J. et al. Peroxisome proliferator-activated receptor gamma coactivator-1 promotes cardiac mitochondrial biogenesis. *J Clin Invest* **106**, 847-56 (2000).

13. Puigserver, P. & Spiegelman, B.M. Peroxisome Proliferator-Activated Receptor-gamma Coactivator 1alpha (PGC-1alpha): Transcriptional Coactivator and Metabolic Regulator. *Endocr Rev* **24**, 78-90 (2003).

14. Qiu, P. Recent advances in computational promoter analysis in understanding the transcriptional regulatory network. *Biochem Biophys Res Commun* **309**, 495-501 (2003).

15. Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J. & Church, G.M. Systematic determination of genetic network architecture. *Nat Genet* **22**, 281-5 (1999).

16. Liu, X.S., Brutlag, D.L. & Liu, J.S. An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nat Biotechnol* **20**, 835-9 (2002).

17. Conlon, E.M., Liu, X.S., Lieb, J.D. & Liu, J.S. Integrating regulatory motif discovery and genome-wide expression analysis. *Proc Natl Acad Sci U S A* **100**, 3339-44 (2003).

18. Bussemaker, H.J., Li, H. & Siggia, E.D. Regulatory element detection using correlation with expression. *Nat Genet* **27**, 167-71 (2001).

19. Zhong, S., Li, C. & Wong, W.H. ChipInfo: Software for extracting gene annotation and gene ontology information for microarray analysis. *Nucleic Acids Res* **31**, 3483-6 (2003).

20. Johnston, S.D. et al. Estrogen-related receptor alpha 1 functionally binds as a monomer to extended half-site sequences including ones contained within estrogen-response elements. *Mol Endocrinol* **11**, 342-52 (1997).

21. Schreiber, S.N., Knutti, D., Brogli, K., Uhlmann, T. & Kralli, A. The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor estrogen-related receptor alpha (ERRalpha). *J Biol Chem* **278**, 9013-8 (2003).

22. Huss, J.M., Kopp, R.P. & Kelly, D.P. Peroxisome proliferator-activated receptor coactivator-1alpha (PGC-1alpha) coactivates the cardiac-enriched nuclear receptors estrogen-related receptor-alpha and -gamma. Identification of novel leucine-rich interaction motif within PGC-1alpha. *J Biol Chem* **277**, 40265-74 (2002).

23. Kamei, Y. et al. PPAR{gamma} coactivator 1{beta}/ERR ligand 1 is an ERR protein ligand, whose expression induces a high-energy expenditure and antagonizes obesity. *Proc Natl Acad Sci U S A* **100**, 12378-83 (2003).

24. Luo, J. et al. Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor alpha. *Mol Cell Biol* **23**, 7947-56 (2003).

25. Chinenov, Y., Coombs, C. & Martin, M.E. Isolation of a bi-directional promoter directing expression of the mouse GABPalpha and ATP synthase coupling factor 6 genes. *Gene* **261**, 311-20 (2000).

26. Batchelor, A.H., Piper, D.E., de la Brousse, F.C., McKnight, S.L. & Wolberger, C. The structure of GABPalpha/beta: an ETS domain- ankyrin repeat heterodimer bound to DNA. *Science* **279**, 1037-41 (1998).

27. Virbasius, J.V., Virbasius, C.A. & Scarpulla, R.C. Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. *Genes Dev* **7**, 380-92 (1993).

28. Virbasius, J.V. & Scarpulla, R.C. Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. *Proc Natl Acad Sci U S A* **91**, 1309-13 (1994).

29. Baar, K. et al. Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1. *FASEB J* **16**, 1879-1886 (2002).

30. Pilegaard, H., Saltin, B. & Neufer, P.D. Exercise induces transient transcriptional activation of the PGC-1alpha gene in human skeletal muscle. *J Physiol* **546**, 851-8 (2003).

31. Davidson, E.H. et al. A genomic regulatory network for development. *Science* **295**, 1669-78 (2002).

32. Sladek, R., Bader, J.A. & Giguere, V. The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. *Mol Cell Biol* **17**, 5400-9 (1997).

33. Herzig, S. et al. CREB regulates hepatic gluconeogenesis through the coactivator PGC-1. *Nature* **413**, 179-83 (2001).

34. Rhee, J. et al. Regulation of hepatic fasting response by PPARgamma coactivator-1alpha (PGC-1): Requirement for hepatocyte nuclear factor 4alpha in gluconeogenesis. *Proc Natl Acad Sci U S A* **100**, 4012-7 (2003).

35. Scarpulla, R.C. Transcriptional activators and coactivators in the nuclear control of mitochondrial function in mammalian cells. *Gene* **286**, 81-9 (2002).

36. Baar, K. et al. Skeletal muscle overexpression of nuclear respiratory factor 1 increases glucose transport capacity. *FASEB J* **17**, 1666-73 (2003).

37. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E.S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241-54 (2003).

38. Cliften, P. et al. Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* **301**, 71-6 (2003).

39. Wasserman, W.W., Palumbo, M., Thompson, W., Fickett, J.W. & Lawrence, C.E. Human-mouse genome comparisons to locate regulatory sites. *Nat Genet* **26**, 225-8 (2000).

40. Ichida, M., Nemoto, S. & Finkel, T. Identification of a specific molecular repressor of the peroxisome proliferator-activated receptor gamma Coactivator-1 alpha (PGC-1alpha). *J Biol Chem* **277**, 50991-5 (2002).

41. Schadt, E.E., Li, C., Ellis, B. & Wong, W.H. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *J Cell Biochem* **Suppl 37**, 120-5 (2001).

42. Mootha, V.K. et al. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115**, 629-40 (2003).

43. Schwartz, S. et al. Human-mouse alignments with BLASTZ. *Genome Res* **13**, 103-7 (2003).

44. Handschin, C., Rhee, J., Lin, J., Tarr, P.T. & Spiegelman, B.M. An autoregulatory loop controls peroxisome proliferator-activated receptor gamma coactivator 1alpha expression in muscle. *Proc Natl Acad Sci U S A* **100**, 7111-6 (2003).

45. St-Pierre, J. et al. Bioenergetic analysis of peroxisome proliferator-activated receptor gamma coactivators 1alpha and 1beta (PGC-1alpha and PGC-1beta) in muscle cells. *J Biol Chem* **278**, 26597-603 (2003).

46. Stunnenberg, H.G. Mechanisms of transactivation by retinoic acid receptors. *Bioessays* **15**, 309-15 (1993).

47. Chinenov, Y., Henzl, M. & Martin, M.E. The alpha and beta subunits of the GA-binding protein form a stable heterodimer in solution. Revised model of heterotetrameric complex assembly. *J Biol Chem* **275**, 7749-56 (2000).

Figure and Table Legends.

Figure 1. Schematic overview of motifADE and application to the PGC-1α timecourse. (a) motifADE attempts to identify motifs associated with differential expression. It begins with a list of genes ordered on the basis of differential expression across two conditions (Condition 1 vs. Condition 2, *e.g.*, healthy vs. disease tissue, treated vs. control, etc.). Each gene is then annotated for the presence or absence of a given motif in the promoter region. A nonparametric statistic is used to assess whether genes with the motif tend to rank high on this list. In this example, genes with Motif 1 are randomly distributed on the list, while genes with Motif 2 tend to rank high. When applied to a single motif or a small set of motifs, the nominal *P*-value associated with the statistic is presented. When we test all possible 6-9-mers, we report the Bonferroni corrected *P*-values. **(b)** C2C12 cells were infected with an adenovirus expressing either GFP (control) or with PGC-1α and profiled over a three day period. Experiments were performed in duplicate and relative gene expression measures are shown. motifADE was applied to the day 3 expression levels for the 5034 genes for which reliable annotations of promoters are available. Genes are ranked according to the difference in expression between PGC-1α and GFP on day 3. Mouse genes having a perfect Errα motif (5'-TGACCTTG-3') or Gabpa/b motif (5'-CTTCCG-3') are labeled with a black bar on the right side of the correlogram. The top 10% of all genes on this rank ordered list contain 25% of all genes with an Errα motif, 14% of all genes with a Gabpa motif, and 34% of all genes containing both motifs. Similarly, the top 25% of all genes on this rank ordered list contain 43% of all genes with an Errα motif, 32% of all genes with the

Gabpa motif, and 52% of all genes with both motifs. **(c)** The expression of 525 nuclear encoded mitochondrial genes, based on a recent genomic and proteomic survey of mouse mitochondria [42], during a three day time-course following infection with PGC-1α or with GFP. Genes were hierarchically clustered using dChip [41] and colors represent relative expression levels.

Figure 2. Proposed model of mechanism of action of PGC-1a. PGC-1α is a highly regulated gene that responds to external stimuli. When PGC-1α levels rise, the expression of Errα and Gabpa are immediately induced via a double positive feedback loop. This results in the strong induction of Errα as well as Gabpa. These levels rise and over the course of 2-3 days, in which these factors couple with PGC-1α to induce the expression of NRF-1 as well as hundreds of effector genes.

Figure 3. Erra and Gabpa cooperate with PGC-1a to induce their own expression. **(a)** Putative Errα and Gabpa motifs 1 kb upstream and downstream of the Errα and the Gabpa transcriptional start sites. **(b), (c), (d)** C2C12 cells were transfected with a reporter gene plasmid containing **(b)** 2 kb of the Errα promoter, **(c)** Gabpa promoter, or **(d)** Gabpa intron 1, together with expression plasmids for Errα, Gabpa, Gabpb1 and PGC-1α. 48 hours post-transfection, reporter gene levels were determined and normalized to β-gal levels.

Figure 4. A small molecule inverse agonist of Erra inhibits the response to PGC-1a. (a) C2C12 myoblasts transfected with wildtype Errα promoter and the respective

expression plasmids were treated with vehicle (0.1% DMSO) or 1 μM XCT790 for 48 hours before reporter gene levels were measured. **(b)** C2C12 cells were transfected with wildtype or Errα-promoter harboring a mutated Errα motif (mutant) together with expression plasmids for Errα, Gabpa, Gabpb1 and PGC-1α. After 48 hours, reporter gene levels were determined and normalized to β-gal levels. **(c)** C2C12 myotubes were infected with GFP- or PGC-1α adenovirus and treated with vehicle (0.1% DMSO) or 1 μM XCT790 for 1 day. Then, relative expression levels of several genes were determined by semiquantitative real-time PCR and normalized against 18S rRNA levels. **(d)** Fao rat hepatoma cells were infected with adenoviral GFP or PGC-1α and treated with vehicle (0.1% DMSO) or 1 μM XCT790 for 1 day before relative gene expression levels were measured by real-time PCR and normalized against 18S rRNA levels. *, P<0.5; ns, not significant.

Figure 5. Total and uncoupled mitochondrial respiration are inhibited by the synthetic Errα-inverse agonist XCT790. (a), (b) C2C12 myotubes were infected with GFP or PGC-1α adenovirus and treated with vehicle (0.1% DMSO) or 1 μM XCT790 for 2 days before total mitochondrial respiration (a) and uncoupled respiration (b) was determined. *, P<0.5 in paired t-test.

Supplemental Figure S1. Cooperativity between the Errα and Gabpa binding sites. All 5034 genes from motifADE analysis are rank ordered on the basis of expression difference (signal:noise metric) on day 3 between cells treated with PGC-1α vs. GFP. The cumulative fraction of genes with a specified motif (Errα, blue; Gabpa, pink; both,

black) is plotted as a function of fractional rank ordering of all 5034 genes. **Figure S2.** Activation of PPARγ does not mimic the effect of the Errα inverse agonist XCT790. C2C12 myotubes were infected with either GFP- or PGC-1α-encoding adenovirus and treated with 1 μM of rosiglitazone for 24 hours before RNA was harvested and relative gene expression determined using real-time PCR. **Figure S3.** PGC-1α-mediated induction of NRF-1 is decreased by the Errα inverse agonist XCT790. C2C12 myotubes were infected with either GFP- or PGC-1α-encoding adenovirus and treated with 1 μM of XCT790 for 24 hours before RNA was harvested and relative gene expression determined using real-time PCR.

Table 1. Motifs associated with differential expression on days 1, 2, and 3. motifADE was performed using the mouse promoter database on each of days 1, 2, and 3. All motifs achieving a Bonferroni-corrected *P*-value < 1 x 10^{-3} are shown. Annotations of the motif and the literature references, when available, are indicated.

Day	Motif	Frequency	Nominal *P*-value	Adjusted *P*-value	Annotation	Reference
1	TGACCTTG	0.07	3.15E-11	2.06E-06	Errα	[20]
	TGACCTTGA	0.02	4.59E-10	1.20E-04	Errα	
2	TGACCTTG	0.07	4.44E-14	2.91E-09	Errα	[20]
	TGACCTT	0.16	3.62E-12	5.93E-08	Errα	
	TGACCT	0.45	1.46E-11	5.97E-08	NR half-site	[46]
	GACCTTG	0.16	7.92E-11	1.30E-06	Errα	
	GACCTT	0.41	1.42E-09	5.81E-06	Errα	
	TTGACC	0.27	2.42E-07	9.92E-04	Errα	
3	CTTCCG	0.33	2.19E-12	8.97E-09	Gabpa	[47]
	TGACCTTG	0.07	1.17E-11	7.66E-07	Errα	[20]
	TGACCTT	0.16	1.23E-10	2.02E-06	Errα	
	CCCGCC	0.54	2.04E-08	8.36E-05		
	GCGGCG	0.43	3.78E-08	1.55E-04		
	AGGTCA	0.42	3.90E-08	1.60E-04	NR half-site	[46]
	CTTCCGG	0.16	1.95E-08	3.19E-04	Gabpa	
	TTCCGG	0.31	1.09E-07	4.46E-04	Gabpa	
	GGGGCG	0.54	1.24E-07	5.08E-04		
	TTCCGCT	0.07	3.30E-08	5.41E-04	Gabpa	
	GCCGGC	0.42	1.57E-07	6.44E-04		
	ACTTCCG	0.09	5.11E-08	8.38E-04	Gabpa	

Supplemental Table S1. Motifs discovered using the promoter database achieving *P*<0.05. motifADE was performed using the mouse promoter database on each of days 1, 2, and 3. Motifs achieving a Bonferroni-corrected *P*-value < 0.05 are shown.

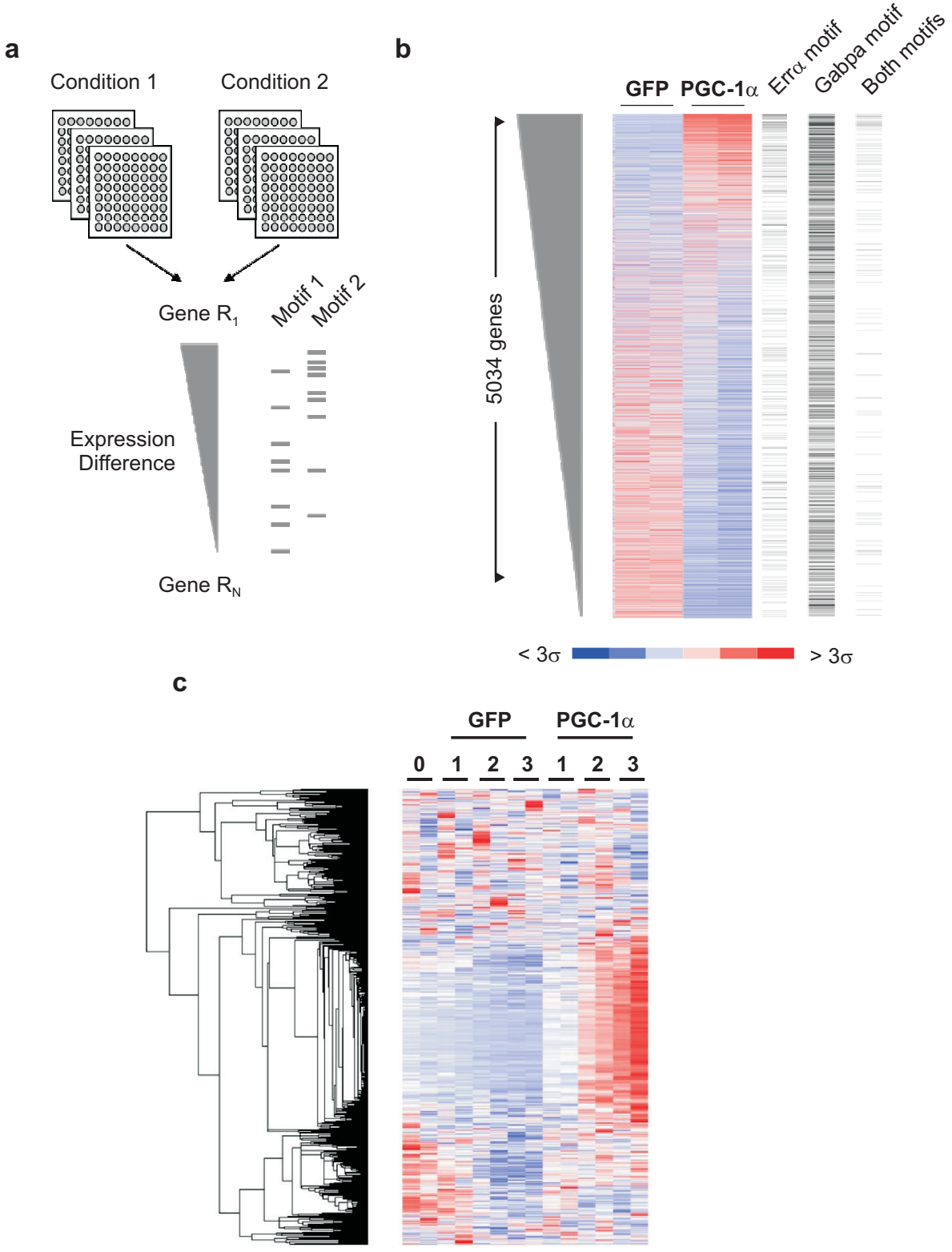
Day	Motif	Frequency	P-value	Adjusted P-value
1	TGACCTTG	0.07	3.15E-11	2.06E-06
	TGACCTTGA	0.02	4.59E-10	1.20E-04
	GACCTTGA	0.05	5.76E-08	3.77E-03
	GACCTTG	0.16	1.54E-06	2.53E-02
	GTCACG	0.18	8.04E-06	3.29E-02
2	TGACCTTG	0.07	4.44E-14	2.91E-09
	TGACCTT	0.16	3.62E-12	5.93E-08
	TGACCT	0.45	1.46E-11	5.97E-08
	GACCTTG	0.16	7.92E-11	1.30E-06
	GACCTT	0.41	1.42E-09	5.81E-06
	TTGACC	0.27	2.42E-07	9.92E-04
	GTGACCTT	0.05	3.86E-08	2.53E-03
	GTGACCT	0.15	3.91E-07	6.41E-03
	GTGACCTTG	0.02	3.97E-08	1.04E-02
	TGACCTTGA	0.02	4.63E-08	1.21E-02
	AGGTCA	0.42	3.46E-06	1.42E-02
	CGCTGAGG	0.04	3.06E-07	2.01E-02
	GACCTTGA	0.05	3.33E-07	2.19E-02
	AGGTCAC	0.13	1.99E-06	3.26E-02
	GTGACC	0.40	8.80E-06	3.61E-02
3	CTTCCG	0.33	2.19E-12	8.97E-09
	TGACCTTG	0.07	1.17E-11	7.66E-07
	TGACCTT	0.16	1.23E-10	2.02E-06
	CCCGCC	0.54	2.04E-08	8.36E-05
	GCGGCG	0.43	3.78E-08	1.55E-04
	AGGTCA	0.42	3.90E-08	1.60E-04
	CTTCCGG	0.16	1.95E-08	3.19E-04
	TTCCGG	0.31	1.09E-07	4.46E-04
	GGGGCG	0.54	1.24E-07	5.08E-04
	TTCCGCT	0.07	3.30E-08	5.41E-04
	GCCGGC	0.42	1.57E-07	6.44E-04
	ACTTCCG	0.09	5.11E-08	8.38E-04
	GACCTT	0.41	2.72E-07	1.11E-03
	CGGGGC	0.51	4.86E-07	1.99E-03
	ATGGCGGC	0.05	4.76E-08	3.12E-03
	GACCTTG	0.16	1.90E-07	3.12E-03
	CTTCCGGC	0.05	7.34E-08	4.81E-03
	ATGGCGG	0.11	3.24E-07	5.31E-03
	AAGATGGCG	0.03	2.07E-08	5.43E-03
	CCGGGG	0.47	1.43E-06	5.85E-03
	GCGGAC	0.24	1.52E-06	6.23E-03
	GGCGGC	0.48	1.55E-06	6.35E-03

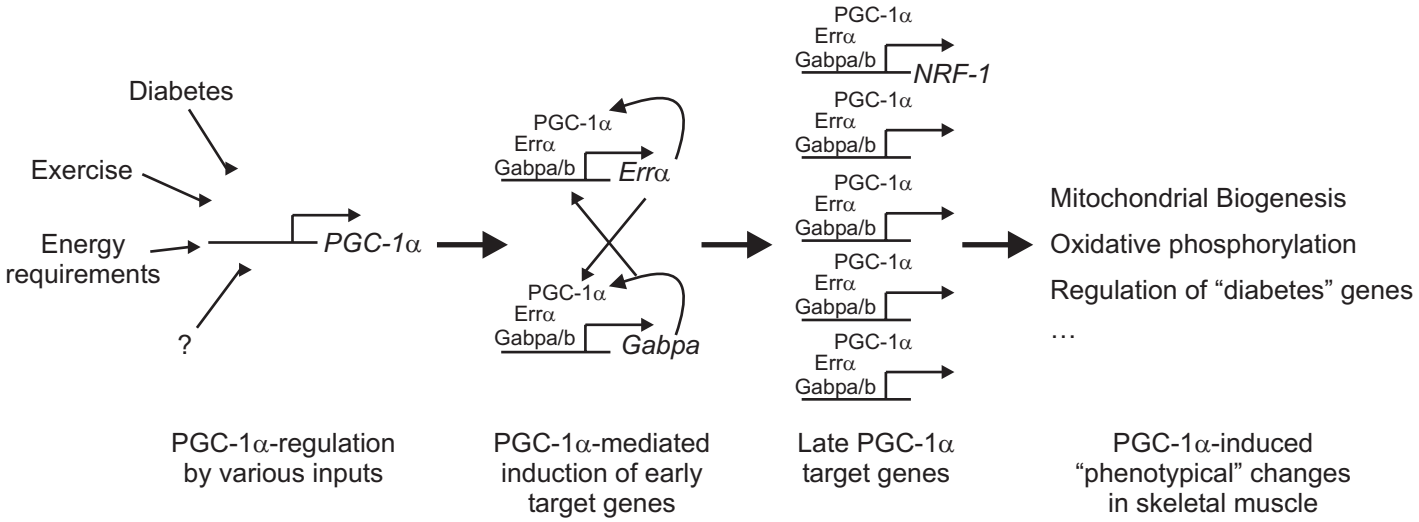
TCACGG	0.19	1.79E-06	7.31E-03
GTGACCTT	0.05	1.23E-07	8.07E-03
CCGGCT	0.39	2.23E-06	9.13E-03
GGCCGG	0.47	2.24E-06	9.16E-03
TCACCG	0.21	2.79E-06	1.14E-02
GCCGGG	0.49	2.81E-06	1.15E-02
CGCCTT	0.30	2.93E-06	1.20E-02
CGGACC	0.24	3.33E-06	1.36E-02
TTCCGC	0.23	3.42E-06	1.40E-02
CGCTGA	0.26	3.44E-06	1.41E-02
CCCCGC	0.51	3.55E-06	1.46E-02
CGCGAG	0.24	3.71E-06	1.52E-02
GTCACG	0.18	4.14E-06	1.69E-02
CGTCCT	0.25	4.15E-06	1.70E-02
AAGGTCA	0.15	1.28E-06	2.10E-02
GCCCGG	0.49	5.14E-06	2.11E-02
CCGCCG	0.36	5.25E-06	2.15E-02
TCCGGG	0.42	5.75E-06	2.35E-02
AAGATGGC	0.08	3.93E-07	2.57E-02
GGCGGA	0.40	6.56E-06	2.69E-02
GGGCGG	0.58	7.63E-06	3.12E-02
CGGGCG	0.38	7.77E-06	3.18E-02
ACCCCG	0.31	8.07E-06	3.30E-02
CGCGCC	0.37	8.13E-06	3.33E-02
CGCCTC	0.41	9.12E-06	3.74E-02
TTCCCG	0.34	9.44E-06	3.86E-02
GGGTCGTGG	0.01	1.56E-07	4.09E-02
CGGCGG	0.40	1.01E-05	4.15E-02
CCGGAA	0.30	1.14E-05	4.68E-02
CGTCGC	0.16	1.15E-05	4.73E-02

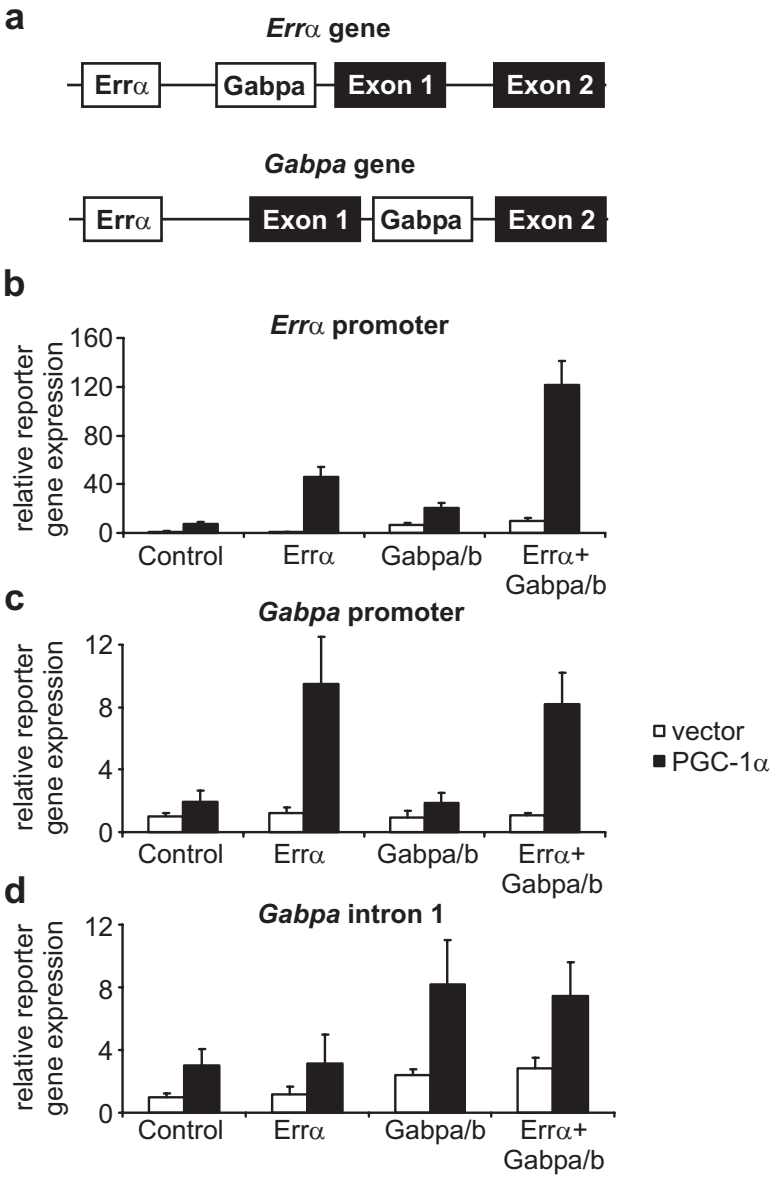
Supplemental Table S2. Motifs discovered using the masked promoter database achieving P<0.05. motifADE was performed using the 'masked promoter database', consisting of regions of the promoters aligned and conserved between mouse and human. Motifs achieving a Bonferroni-corrected *P*-value < 0.05 are shown.

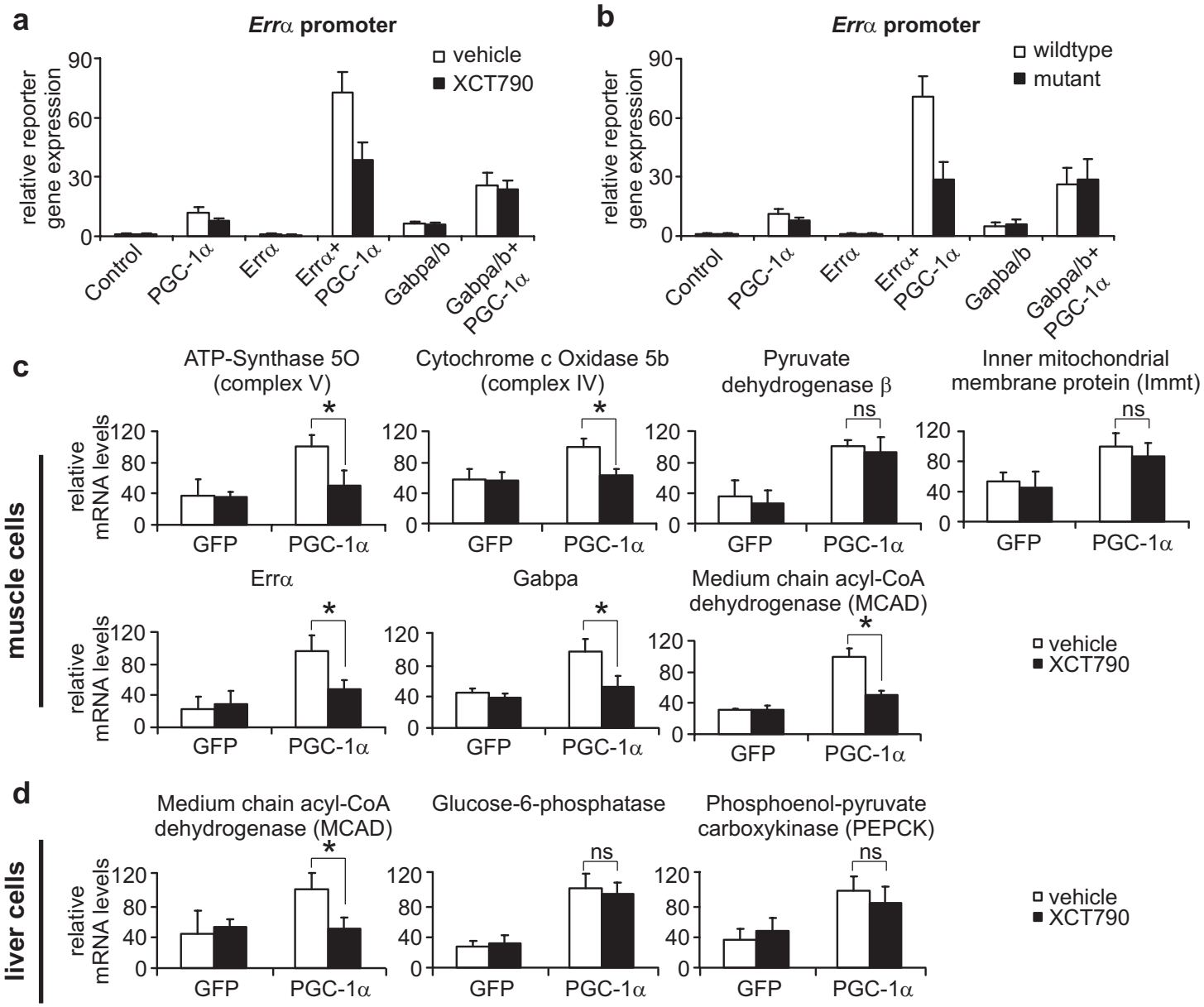
Day	Motif	Frequency	P-value	Adjusted P-value
1	TGACCTTG	0.04	9.93E-12	6.51E-07
	TGACCTT	0.09	2.10E-09	3.44E-05
	GACCTTGA	0.02	4.91E-08	3.22E-03
	GTGACCTTG	0.02	4.30E-08	1.13E-02
	TGACCTTGA	0.01	4.92E-08	1.29E-02
2	TGACCTT	0.09	1.09E-14	1.78E-10
	TGACCT	0.26	4.56E-13	1.87E-09
	GTGACCTT	0.03	6.37E-13	4.18E-08
	TGACCTTG	0.04	8.30E-11	5.44E-06
	GTGACCT	0.09	5.72E-10	9.37E-06
	AGGTCA	0.25	2.90E-09	1.19E-05
	GACCTT	0.24	8.67E-09	3.55E-05
	GACCTTG	0.09	4.78E-08	7.84E-04
	ACCTTG	0.22	1.06E-06	4.33E-03
	TGTGACCTT	0.01	6.13E-08	1.61E-02
	GTGACCTTG	0.02	1.14E-07	2.98E-02
	AGGTCAC	0.08	2.01E-06	3.29E-02
	ACCTTGA	0.06	2.63E-06	4.30E-02
3	TGACCTT	0.09	4.40E-14	7.20E-10
	TGACCTTG	0.04	6.24E-12	4.09E-07
	GTGACCTT	0.03	1.37E-11	8.97E-07
	GTGACCT	0.09	7.06E-11	1.16E-06
	CTTCCG	0.26	1.28E-08	5.26E-05
	TGACCT	0.26	2.11E-08	8.64E-05
	ATGGCGGC	0.05	2.87E-09	1.88E-04
	GACCTT	0.24	8.73E-08	3.57E-04
	AGGTCA	0.25	9.47E-08	3.88E-04
	AAGATGGCG	0.03	1.70E-09	4.45E-04
	CGGTGA	0.20	1.81E-07	7.41E-04
	AGATGGCG	0.04	1.31E-08	8.61E-04
	CTTCCGG	0.12	1.66E-07	2.72E-03
	GACCTTG	0.09	2.66E-07	4.35E-03
	AGATGGCGG	0.02	2.22E-08	5.81E-03
	ATGGCGG	0.10	4.34E-07	7.11E-03
	CCGGGG	0.39	1.80E-06	7.39E-03
	GGCGGG	0.53	2.33E-06	9.56E-03
	ACTTCCG	0.08	1.19E-06	1.94E-02
	CGGAAGT	0.08	1.24E-06	2.04E-02

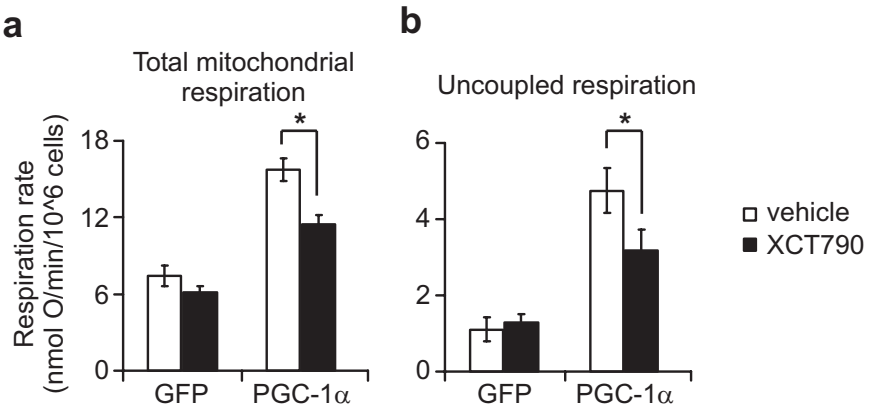
CTTCCGGC	0.04	3.20E-07	2.10E-02
TGGCGGC	0.16	1.70E-06	2.78E-02
CCGGCT	0.28	9.09E-06	3.72E-02
TGACCTTGA	0.01	1.50E-07	3.93E-02
AAGATGGCGG	0.02	4.08E-08	4.28E-02
GAGGTCA	0.08	2.68E-06	4.40E-02
GCGGAA	0.18	1.15E-05	4.73E-02


a
Condition 1
Condition 2
Gene R1
Motif 1
Motif 2
Expression Difference
Gene RN
b
GFP
PGC-1α
Errα motif
Gabpa motif
Both motifs
5034 genes
< 3σ
> 3σ
c
GFP
PGC-1α
0
1
2
3
1
2
3

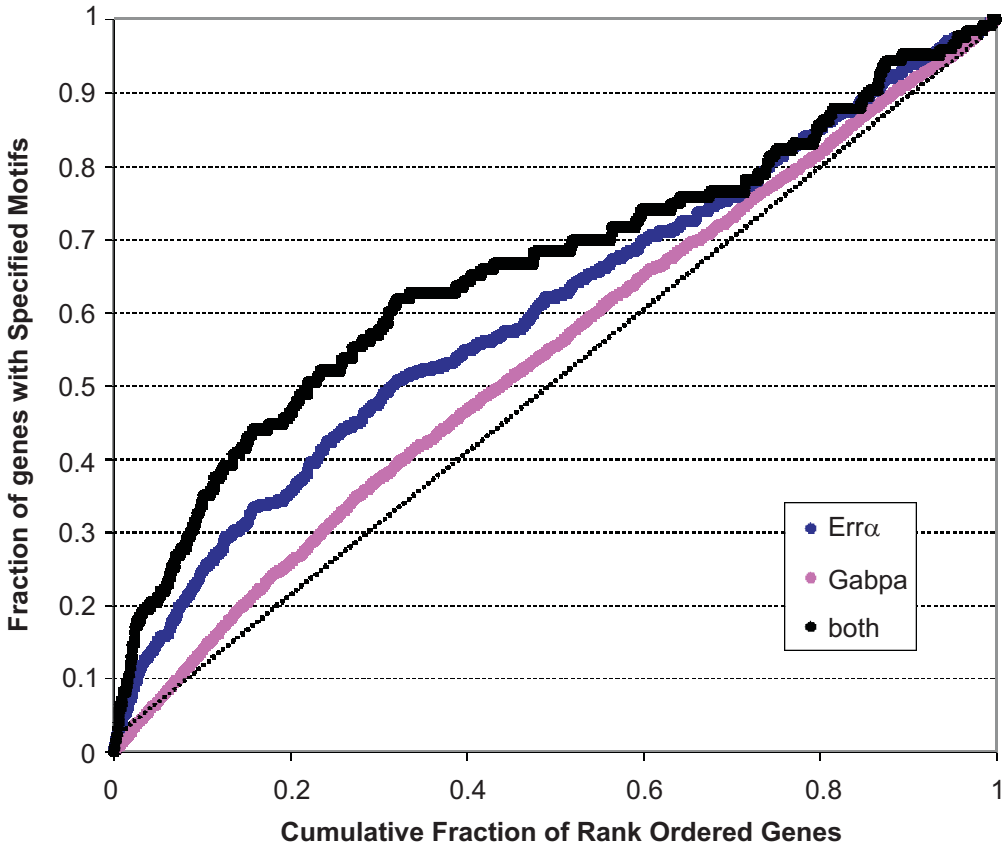

Diabetes
Exercise
Energy requirements
?
PGC-1α
PGC-1α
Errα
Gabpa/b
Errα
PGC-1α
Errα
Gabpa/b
Gabpa
PGC-1α
Errα
Gabpa/b
NRF-1
PGC-1α
Errα
Gabpa/b
PGC-1α
Errα
Gabpa/b
PGC-1α
Errα
Gabpa/b
PGC-1α
Errα
Gabpa/b
Mitochondrial Biogenesis
Oxidative phosphorylation
Regulation of “diabetes” genes
…
PGC-1α-regulation by various inputs
PGC-1α-mediated induction of early target genes
Late PGC-1α target genes
PGC-1α-induced “phenotypical” changes in skeletal muscle


a
Errα gene
Errα
Gabpa
Exon 1
Exon 2
Gabpa gene
Errα
Exon 1
Gabpa
Exon 2
b
Errα promoter
relative reporter gene expression
160
120
80
40
0
Control
Errα
Gabpa/b
Errα+ Gabpa/b
c
Gabpa promoter
relative reporter gene expression
12
8
4
0
Control
Errα
Gabpa/b
Errα+ Gabpa/b
vector
PGC-1α
d
Gabpa intron 1
relative reporter gene expression
12
8
4
0
Control
Errα
Gabpa/b
Errα+ Gabpa/b

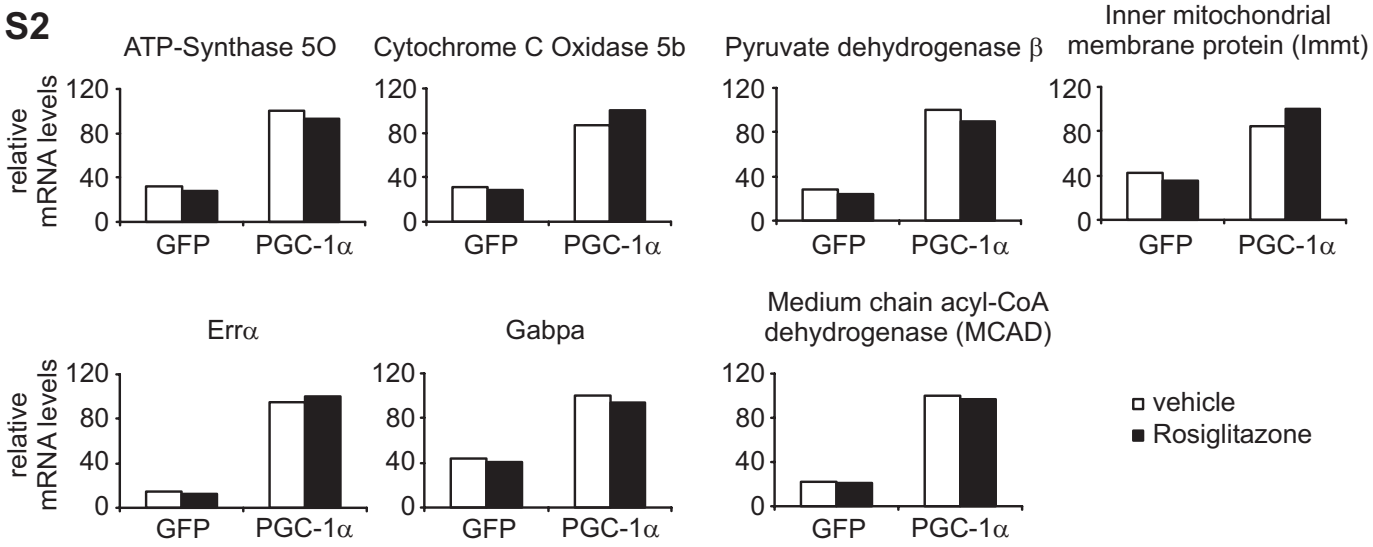

a
Errα promoter
relative reporter gene expression
vehicle
XCT790
Control
PGC-1α
Errα
Errα+ PGC-1α
Gabpa/b
Gabpa/b+ PGC-1α
b
Errα promoter
wildtype
mutant
c
muscle cells
ATP-Synthase 5O (complex V)
Cytochrome c Oxidase 5b (complex IV)
Pyruvate dehydrogenase β
Inner mitochondrial membrane protein (Immt)
Errα
Gabpa
Medium chain acyl-CoA dehydrogenase (MCAD)
relative mRNA levels
GFP
PGC-1α
ns
d
liver cells
Medium chain acyl-CoA dehydrogenase (MCAD)
Glucose-6-phosphatase
Phosphoenol-pyruvate carboxykinase (PEPCK)


a
Total mitochondrial respiration
Respiration rate (nmol O/min/10^6 cells)
18
12
6
0
*
GFP
PGC-1α
b
Uncoupled respiration
6
4
2
0
*
GFP
PGC-1α
vehicle
XCT790

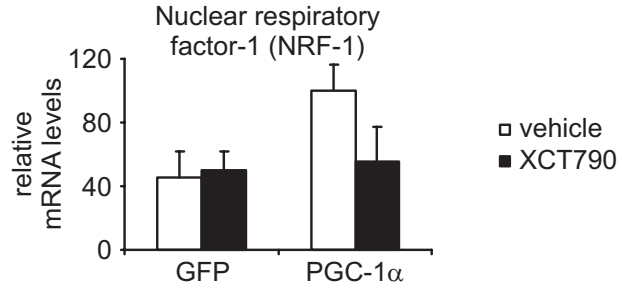
S1


Fraction of genes with Specified Motifs
Cumulative Fraction of Rank Ordered Genes
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.2
0.4
0.6
0.8
Errα
Gabpa
both

S2


ATP-Synthase 5O
Cytochrome C Oxidase 5b
Pyruvate dehydrogenase β
Inner mitochondrial membrane protein (Immt)
Errα
Gabpa
Medium chain acyl-CoA dehydrogenase (MCAD)
relative mRNA levels
120
80
40
0
GFP
PGC-1α
vehicle
Rosiglitazone

S3


Nuclear respiratory factor-1 (NRF-1)
relative mRNA levels
120
80
40
0
GFP
PGC-1α
vehicle
XCT790